

Internet Electronic Journal*

Nanociencia et Moletrónica

Mayo 2007, Vol. 5, N°1, pp. 921-940

Estudio bioinformático de la molécula proteica UgpB de bacterias del género *Brucella*

L. Morales^{1,2}, A. Mena², E. I. Castañeda¹, V. Serkin³, R. Peña³

Benemérita Universidad Autónoma de Puebla,
¹ Centro de Investigaciones en Ciencias Microbiológicas
Ciudad Universitaria, Puebla, Pue., CP 72570, México.

² Facultad de Ciencias Químicas. BUAP.

³ Centro de Química-ICUAP

e-mail: lamoral@siu.buap.mx

recibido: 03. 04. 07

revisado: 15.04.07

publicado: 31.05.07

Citation of the article;

L. Morales, A. Mena, E. I. Castañeda, V. Serkin, R. Peña, Estudio bioinformático de la molécula proteica UgpB de bacterias del género *Brucella*, Internet Electron. J. Nanoc. Moletrón. 2007, Vol. 5, N° 1, pp 921-940

copyright ©BUAP 2007

Estudio bioinformático de la molécula proteica UgpB de bacterias del género *Brucella*

L. Morales^{1,2}, A. Mena², E. I. Castañeda¹, V. Serkin³, R. Peña³

Benemérita Universidad Autónoma de Puebla,
¹ Centro de Investigaciones en Ciencias Microbiológicas
Ciudad Universitaria, Puebla, Pue., CP 72570, México.

² Facultad de Ciencias Químicas. BUAP.

³ Centro de Química-ICUAP
e-mail: lamoral@siu.buap.mx

recibido: 03.04.07

revisado: 15.04.07

publicado: 31.05.07

Internet Electron. J. Nanoc. Moletrón., Vol. 5, N° 1, pp. 921-940

RESUMEN

En este trabajo se determinaron las características de la secuencia de la proteína UgpB de las bacterias *B. melitensis*, *B. suis* y *B. abortus*, con respecto a su homología con proteínas del mismo género y de otros microorganismos, su probable topología, estructura secundaria, terciaria y su relación funcional mediante herramientas de bioinformática, empleando programas computacionales de predicción y búsqueda de secuencias en los bancos de datos. Se encontró que UgpB es una proteína transmembranal de tipo II, cuyo dominio extracelular podría reconocer azúcares como trehalosa, que puede funcionar como un osmoprotector bacteriano para soportar altas concentraciones salinas, semejante a lo que sucede en microorganismos que habitan en aguas saladas de océanos. Otro azúcar que podría interaccionar con este dominio es el ácido siálico, presente en las membranas de las células de los organismos hospederos de la bacteria, por lo que este reconocimiento puede ser fundamental en el desarrollo de los mecanismos de adherencia e invasión de *Brucella* a las células huésped para generar el proceso infeccioso y provocar la enfermedad. Los resultados obtenidos mediante bioinformática fueron muy valiosos, pues, además de estar relacionados con evidencia experimental reportada, nos dan un panorama más amplio de las posibles estrategias que puede emplear la bacteria para soportar diversas condiciones del medio en el que se encuentren, con la posibilidad de infectar a sus huéspedes. Las investigaciones acerca de estos mecanismos moleculares, abren la posibilidad de diseñar mediante nanobiotecnología, nanoestructuras que realicen trabajos específicos y selectivos, con el fin de inhibir mecanismos de patogenicidad inducidos por bacterias ó microorganismos que expresen proteínas homólogas.

Palabras clave: Bioinformática, nanobiotecnología, proteína UgpB, *Brucella*, predicción.

1. Introducción

La determinación de las funciones de una proteína, así como los mecanismos desencadenados para llevar a cabo tales funciones, implica años de experimentación así como una fuerte inversión económica, es por eso, que el empleo de programas de predicción y de comparación para el análisis de estructuras primarias, secundarias y terciarias de las proteínas, brinda la posibilidad de conocer sus funciones más probables al compararlas con otras biomoléculas proteicas homólogas cuyas funciones han sido descritas con gran detalle, lo que permite orientar con gran precisión los protocolos de experimentación antes y después de su realización para comprender con mayor visión los mecanismos por medio de los cuales una biomolécula puede realizar sus actividades ya sean catalíticas, de transporte, mensajeras, de reconocimiento del medio, etc. [1, 2].

La estructura primaria, secundaria y tridimensional que presentan las proteínas, ha sufrido evolución a lo largo de la vida con el objetivo de adquirir la conformación tridimensional más adecuada con la cual la proteína pueda realizar funciones especializadas y así, brindarles a las células a las cuales pertenecen una mejor oportunidad de sobrevivir, adaptarse, desarrollarse, multiplicarse y preservar su especie.

Las bacterias han mostrado tener una gran capacidad de adaptación al medio ambiente, mediante la evolución de sus proteínas, permitiéndoles habitar en distintos ambientes.

Las bacterias del género *Brucella*, son un buen ejemplo de la adaptación a diferentes ambientes, pues se han aislado cepas en ecosistemas como agua, suelos, alimentos, animales y humanos [3]. Las especies del género *Brucella*, son microorganismos gram negativos intracelulares facultativos, que provocan la brucelosis que es una enfermedad ambiental y re-emergente considerada por la Organización de las Naciones Unidas (ONU) como la zoonosis de mayor distribución mundial, la cual además de su efecto patogénico en el hombre tiene un enorme impacto en la industria animal provocando pérdidas económicas a los países [4]. Sin embargo, es una zoonosis desatendida en los países pobres ó en vías de desarrollo, a pesar de provocar daños significativos a la salud humana, pues es una enfermedad severamente debilitante y generalmente prolongada, se caracteriza por provocar fiebre recurrente con sudoración nocturna, dolor de cabeza, escalofrío, malestar general, fatiga, artralgias, depresión, anorexia, con posibilidad de desarrollar neurobrucelosis y granulomas intracerebrales [5, 6, 7].

Las bacterias del género *Brucella* expresan un conjunto de proteínas de superficie, de las cuales Castaneda et al., en el 2004 identifican a SP-41 (Proteína de Superficie de 41 kDa), que es producto del gen UgpB, el cual codifica para la proteína UgpB (Uptake glycerol phosphate). Sin embargo, no existen datos experimentales que confirmen esta función.

La biomolécula UgpB ha sido identificada en cepas aisladas de leches de cabras y vacas con brucelosis, suelos y agua contaminados con la bacteria, así como de hemocultivos humanos de pacientes con brucelosis aguda. La evidencia experimental ha mostrado que UgpB tiene propiedades inmuno-reactivas y como adhesina, importantes en la adherencia a las células epiteliales humanas (células HeLa) mediante la interacción con residuos de ácido siálico presentes en la superficie de éstas células [8].

Los estudios de inmunofluorescencia e inmuno-oro [9] han sugerido que parte de la proteína UgpB puede localizarse sobre la superficie celular bacteriana, sin embargo, se desconoce su estructura secundaria y terciaria, así como su topología y homología con otras proteínas de otros microorganismos.

Por lo que en este trabajo nos enfocamos en la determinación de estas características estructurales y su probable relación funcional, mediante estudios de análisis de predicción

de estructuras, con el objetivo de identificar los dominios de la proteína y los aminoácidos que pudieran interactuar en el reconocimiento del ácido siálico, importante para el desarrollo de la invasión y patogénesis de *Brucella* en su célula huésped.

2. Metodología

Las secuencias empleadas de UgpB de *Brucella melitensis*, *B. abortus* y *B. suis* se obtuvieron de la base de datos del proyecto de la Enciclopedia de Kyoto de genes y genomas (KEGG, www.genome.ad.jp/kegg/) [10], cuyos números de acceso son; BMEI10625, BruAb2_0571 y BRA0655 respectivamente. La búsqueda de homólogos a UgpB se realizó con el programa BLAST (Herramienta de Búsqueda y Alineamiento Local Básico) (<http://www.ncbi.nlm.nih.gov/BLAST/>), [11], empleando la base de secuencias no redundante de proteínas del GenBank, conformada por el Centro Nacional para la Información en Biotecnología (NCBI), el cual contiene todas las secuencias proteicas publicadas hasta la fecha (<http://www.ncbi.nlm.nih.gov/>), [12]. Se generó el alineamiento múltiple de la secuencia de UgpB con las secuencias proteicas de la base de datos mediante el programa ClustalW (<http://www.ebi.ac.uk/clustalw/>) para obtener el alineamiento entre la secuencia de UgpB y sus homologas [13].

La búsqueda de los sitios biológicamente significativos se realizó con el programa Scan Prosite [14], con el programa CD-Search

(<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>) [15] se identificaron dominios conservados en la secuencia de UgpB, este programa emplea modelos basados en alineamientos múltiples de secuencias, con significancia estadística, la base de datos empleada fue PSSMs que genera matrices a las cuales se les asigna un puntaje a los modelos de los dominios obtenidos por alineamiento [16, 17].

La predicción de regiones transmembrana y su topología se obtuvieron mediante programas disponibles en el servidor ExPASy (Expert Protein Analysis System <http://www.expasy.org>) [18], TMHMM (<http://www.cbs.dtu.dk/services/TMHMM-2.0/>) predice presencia, orientación y el tipo de estructura secundaria en la zona transmembranal de la proteína, por medio de modelos matemáticos sofisticados [19, 20, 21], TopPred y ProtScale, (<http://bioweb.pasteur.fr/seqanal/interfaces/toppred.html>) predicen la topología mediante el análisis de hidrofobicidad de los aminoácidos que integran su secuencia [22].

Pfam (<http://www.sanger.ac.uk/Software/Pfam/index.shtml>) se empleó para para conocer en que familia de proteínas clasifica a UgpB de acuerdo sus posibles dominios y predijo regiones transmembranales.

Para la predicción de estructura secundaria se empleó el programa PHD, (<http://cubic.bioc.columbia.edu/predictprotein/>) [23], así como el servidor PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred/>) [24]. La estructura terciaria de la proteína fue modelada mediante el programa SWISS-MODEL (<http://swissmodel.expasy.org/>) que es un servidor completamente automatizado para el modelado por homología de estructuras de proteínas [25], para la visualización de los modelos generados y la superposición de las estructuras terciarias se empleó el programa Swiss-PdbViewer [26].

3. Resultados y Discusión

3.1. Análisis de la secuencia primaria de la biomolécula protéica UgpB de bacterias del género *Brucella*.

Las secuencias proteicas UgpB de *Brucella melitensis*, *B. abortus* y *B. suis* fueron analizadas mediante bioinformática, en su estructura primaria y probables características en su estructura secundaria y terciaria (tridimensional), empleando programas de predicción, con el objetivo de tener un panorama más amplio acerca de su estructura y relación funcional.

Inicialmente se llevó a cabo el análisis de la similitud de la secuencia de la proteína UgpB con respecto a secuencias reportadas en el banco de datos de secuencias de proteínas, en búsqueda de posibles relaciones funcionales, puesto que secuencias similares, es decir, homólogas, es muy probable que tengan el mismo ancestro y compartan características en su estructura y función biológica.

3.1.1. Búsqueda de homólogos a la biomolécula proteica UgpB.

La búsqueda de homólogos mediante BLAST mostró (Tabla 1) que las tres secuencias UgpB de *Brucella* tienen una homología del 99%, UgpB de *B. melitensis* es altamente homóloga a UgpB de los microorganismos *Rhizobium meliloti* (*Sinorhizobium meliloti*) 85%, *Rhodobacter spheroides* 71%, *Mesorhizobium loti* 37%, *Agrobacterium tumefaciens* 37%, estos microorganismos son miembros de la subdivisión del grupo de las α -proteobacterias, que son patógenos o simbioses intracelulares de plantas, relacionados filogenéticamente entre sí, y con los Gram negativos *Escherichia coli* y *Salmonella enterica* presenta una homología del 48 y 46% respectivamente, esta homología se considera alta [2] y el valor -E (valor de expectación) para cada una de ellas indica que el resultado de la similitud de la secuencia analizada (UgpB de *B. melitensis*) y las secuencias homólogas encontradas en la base de datos tiene significación estadística con valores -E de 0 ó muy pequeños (menores de 10^{-4}) [11, 27].

Tabla 1. HOMOLOGÍA ENTRE LAS SECUENCIAS PROTEICAS DE UgpB DE *B. melitensis* CON OTROS ORGANISMOS

Organismo	Proteína	Homología	Valor-E
<i>B. melitensis</i>	UgpB	100	
<i>B. suis</i>	UgpB	99	0
<i>B. abortus</i>	UgpB	99	0
Rhizobium meliloti (Sinorhizobium meliloti) *	UgpB	85	0
<i>Rhodobacter spheroides</i> *	UgpB	71	4e-177
<i>E. Coli</i>	UgpB	48	0
<i>Salmonella enterica</i>	UgpB	46	5e-110
<i>Mesorhizobium loti</i> *	UgpB	37	4e-75
<i>Agrobacterium tumefaciens</i> *	UgpB	37	0

* Microorganismos del grupo de las α -Proteobacterias

3.1.2. Diferencia en las secuencias UgpB en el género *Brucella*.

Debido a que se encontró que las tres secuencias proteicas UgpB reportadas en el banco de datos para *B. suis*, *B. melitensis* y *B. abortus*, tienen una homología del 99 %, examinamos la diferencia entre estas 3 secuencias.

Las diferencias entre las secuencias de UgpB de las 3 especies de *Brucella*, se identificaron en el alineamiento ClustalW (Fig 1), la primera diferencia consiste en que la proteína UgpB de *B. melitensis* tiene 444 residuos aminoacídicos, cuyos primeros 11 aminoácidos (aa) no están presentes en las secuencias de *B. abortus* y *B. suis*, por lo que UgpB en estas especies solo tienen 433 aa, hasta la fecha no hay evidencia experimental de que el precursor UgpB de *B. melitensis* sea cortado para funcionar como una proteína madura de 433 residuos, dado que tiene dos metioninas de inicio. La segunda diferencia, está en el residuo 10 que en es alanina en *B. suis* y *B. melitensis* y valina en *B. abortus*. Y la tercera diferencia es que *B. abortus* y *B. suis* tienen treonina en la posición 334 y *B. melitensis* tiene serina, no obstante, los cambios de aa en estas posiciones pudieran no inducir un efecto prominente en cuanto a estructura y función de UgpB, pues las características fisicoquímicas de ellos son muy similares.

```

tr|Q578E6|Q578E6_BRUAB -----MFTRLITTSVLTGAIALTIGSQAFQAQTELAWWHGTMGAN 39
tr|Q8FW10|Q8FW10_BRUSU -----MFTRLITTSALTGAIALTIGSQAFQAQTELAWWHGTMGAN 39
tr|Q8YCA7|Q8YCA7_BRUME MFTQSYIRGGNMFTRLITTSALTGAIALTIGSQAFQAQTELAWWHGTMGAN 50
      *****
tr|Q578E6|Q578E6_BRUAB NEMVNLSKEFNESQSEYKIVPVYKGNYPETLNAGIAAFRSKQPPAILQV 89
tr|Q8FW10|Q8FW10_BRUSU NEMVNLSKEFNESQSEYKIVPVYKGNYPETLNAGIAAFRSKQPPAILQV 89
tr|Q8YCA7|Q8YCA7_BRUME NEMVNLSKEFNESQSEYKIVPVYKGNYPETLNAGIAAFRSKQPPAILQV 100
      *****
tr|Q578E6|Q578E6_BRUAB FDAGSGVMMAAEGAIVPAAEVLEKGGYKFDKSQYLPGIVAYYSKPDGTM 139
tr|Q8FW10|Q8FW10_BRUSU FDAGSGVMMAAEGAIVPAAEVLEKGGYKFDKSQYLPGIVAYYSKPDGTM 139
tr|Q8YCA7|Q8YCA7_BRUME FDAGSGVMMAAEGAIVPAAEVLEKGGYKFDKSQYLPGIVAYYSKPDGTM 150
      *****
tr|Q578E6|Q578E6_BRUAB SFPYNSSSPILYYNKDAFKKAGLDENKPPKTWPEVFEEAAKKIKASGASPC 189
tr|Q8FW10|Q8FW10_BRUSU SFPYNSSSPILYYNKDAFKKAGLDENKPPKTWPEVFEEAAKKIKASGASPC 189
tr|Q8YCA7|Q8YCA7_BRUME SFPYNSSSPILYYNKDAFKKAGLDENKPPKTWPEVFEEAAKKIKASGASPC 200
      *****
tr|Q578E6|Q578E6_BRUAB GFTSTWLTWIQTENFAAWNVPYGTNENGLAGTDVKLEINSPLYVEHFQA 239
tr|Q8FW10|Q8FW10_BRUSU GFTSTWLTWIQTENFAAWNVPYGTNENGLAGTDVKLEINSPLYVEHFQA 239
tr|Q8YCA7|Q8YCA7_BRUME GFTSTWLTWIQTENFAAWNVPYGTNENGLAGTDVKLEINSPLYVEHFQA 250
      *****
tr|Q578E6|Q578E6_BRUAB IADLAKDGTFRYGGRTSEAKQLFTSGECAMLTSSGGLGDVVKSGINYG 289
tr|Q8FW10|Q8FW10_BRUSU IADLAKDGTFRYGGRTSEAKQLFTSGECAMLTSSGGLGDVVKSGINYG 289
tr|Q8YCA7|Q8YCA7_BRUME IADLAKDGTFRYGGRTSEAKQLFTSGECAMLTSSGGLGDVVKSGINYG 300
      *****
tr|Q578E6|Q578E6_BRUAB GQLPYYEGHGPQNTIPGGASLWVFAGLSDDQYKGIAEFFNFLSQTKIQVK 339
tr|Q8FW10|Q8FW10_BRUSU GQLPYYEGHGPQNTIPGGASLWVFAGLSDDQYKGIAEFFNFLSQTKIQVK 339
tr|Q8YCA7|Q8YCA7_BRUME GQLPYYEGHGPQNTIPGGASLWVFAGLSDDQYKGIAEFFNFLSQTKIQVK 350
      *****
tr|Q578E6|Q578E6_BRUAB LHEKSGYLPVTLAAYEETKKSDFYEKNPGRETPIQLQMMGKEPTENSKGVR 389
tr|Q8FW10|Q8FW10_BRUSU LHEKSGYLPVTLAAYEETKKSDFYEKNPGRETPIQLQMMGKEPTENSKGVR 389
tr|Q8YCA7|Q8YCA7_BRUME LHEKSGYLPVTLAAYEETKKSDFYEKNPGRETPIQLQMMGKEPTENSKGVR 400
      *****
tr|Q578E6|Q578E6_BRUAB LVNLPQVRDILNEEFAEMLGGKQDAKTALDNAVKRGNAIAAAQ 433
tr|Q8FW10|Q8FW10_BRUSU LVNLPQVRDILNEEFAEMLGGKQDAKTALDNAVKRGNAIAAAQ 433
tr|Q8YCA7|Q8YCA7_BRUME LVNLPQVRDILNEEFAEMLGGKQDAKTALDNAVKRGNAIAAAQ 444
      *****

```

Fig. 1. Alineamiento de secuencias de *B. melitensis* (BRUME), *B. suis* (BRUSU), *B. abortus* (BRUAB) mediante CLUSTAL W (1.83). *B. melitensis*, 444 aa; *B. suis*, 433 aa; *B. abortus*, 433 aa. Las 3 diferencias en las secuencias están señaladas con color amarillo, la identidad de secuencia se señala con un asterisco (*).

3.1.3. Búsqueda de un motivo de unión a sustrato en *B. melitensis* y comparación del dominio en microorganismos homólogos

La homología entre proteínas se encuentra más conservada en estructura que en secuencia aminoacídica y no es raro encontrar en la naturaleza que diferentes proteínas presenten dominios conservados, aún cuando su identidad de secuencia sea baja, conservándose así funciones básicamente muy similares, incluso más específicas, debido a que los dominios pueden sufrir modificación durante la evolución para adecuarse a una nueva función como sería la especialización por la afinidad de unión a un sustrato determinado en función de los requerimientos señalados por las condiciones ambientales en las cuales el organismo o microorganismo se ha adaptado con mayor eficacia y ha evolucionado desarrollando funciones especializadas. Se sabe que es probable que existan no más de pocos miles de diferentes tipos de dominios. Sin embargo, la naturaleza construye nuevas proteínas con esta colección, caracterizada por la estabilidad de sus formas [2]. Por lo que resulta interesante realizar la búsqueda de sitios o dominios conservados que pudieran localizarse en una proteína de función desconocida y así conocer e identificar su posible interacción con una proteína, sustrato, metal, etc.

Con el programa Scan Prosite y con el servidor de Dominios Conservados, se identificó un sitio de unión a soluto extracelular bacteriano con la secuencia consenso PFNSSTPALFYNKDLFKKAGL del residuo 143-157 de UgpB (Fig 2A), el alineamiento de las secuencias tiene un valor -E de $1e-49$ (CD-search), que es de alta confiabilidad, este dominio está conservado (Fig 2B) en las proteínas UgpB de las 8 secuencias homólogas mostradas en la tabla 1, todas estas proteínas tiene una función putativa asignada como proteínas de alta afinidad de unión a soluto extracelular, en particular se les ha denominado proteínas de unión a glicerol-fosfato, las cuales tienen la características de formar parte de sistemas de transportadores de tipo ABC [28], presentes en las membranas celulares tanto de procariontes como de eucariontes unicelulares y multicelulares, con afinidades muy particulares por la unión de sustratos, algunos transportadores expulsan ciertas moléculas, otros las introducen del medio exterior, otros actúan en ambas direcciones. No obstante, estas funciones para UgpB en el género *Brucella*, no han sido comprobadas experimentalmente pues aún se desconoce mucho acerca de su función.



Fig. 2. Sitio de unión a soluto extracelular en UgpB. A. Servidor CD, secuencia consenso del motivo de unión a soluto extracelular (SBP) en *Brucella abortus* de Prolina 143 a Leucina 163 subrayado en color rojo. **B.** Localización del motivo SBP en la secuencia UgpB de los 8 microorganismos homólogos, mediante CLUSTALW, el motivo se encuentra conservado, los colores de los aminoácidos indican su carácter polar a pH=7, aa ácidos con carga positiva en rosa, aa ácidos con carga negativa en azul, aa polares en negro, no polares en rojo.

3.2. Análisis de la topología y estructura secundaria de UgpB

3.2.1. Predicción de segmentos transmembranales

El análisis de la secuencia primaria de UgpB nos permitió determinar mediante programas de predicción, la topología de la proteína y su probable estructura secundaria. Este análisis mostró la presencia de un segmento transmembranal en el extremo amino, orientado hacia el interior de la célula y el resto de la proteína fue ubicado hacia el exterior celular, lo cual se corroboró mediante diferentes programas. En la Figura 3 mostramos el gráfico obtenido mediante TMHMM, que predice la presencia de ~12 aa del extremo amino en el interior celular y del residuo 13 al 35 se localiza el segmento transmembranal,

mientras el resto de la proteína alrededor del residuo 36 en adelante es ubicado hacia el exterior de la membrana.

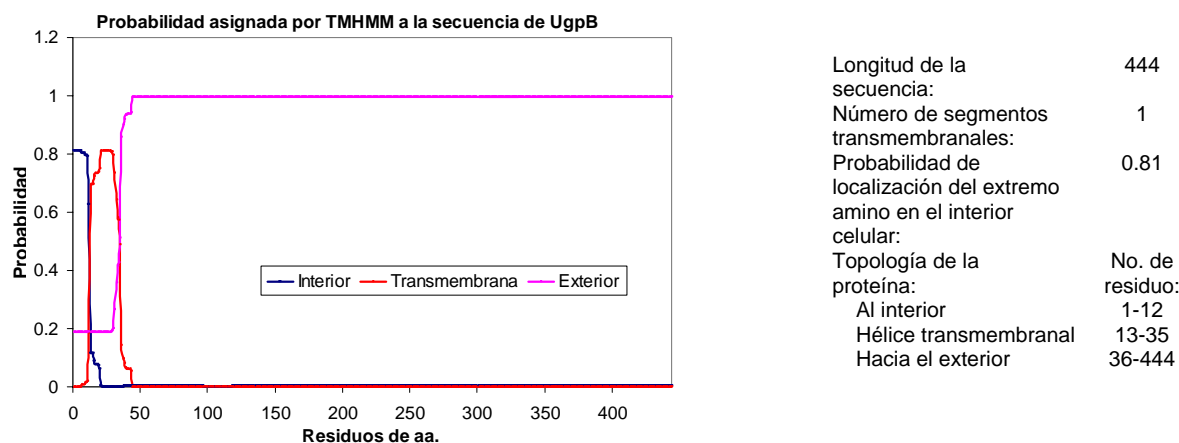


Fig. 3 Predicción de segmentos transmembranales y topología de UgpB.

Transmembrana —, Interior —, Exterior —.

TopPred confirmó el resultado obtenido mediante TMHMM, el gráfico de la Fig 4, mostró un pico con el mayor puntaje de hidrofobicidad localizado en el extremo amino de la proteína, presentando el mismo patrón que los resultados obtenidos mediante ProtScale, estos resultados también fueron corroborados con las escalas de Kyte-Doolittle [29], Eisenberg [30], Rao y Argos [31], empleando ventana de 19 [2], estos programas mostraron gráficos similares para la predicción de la región transmembranal de UgpB (gráficos no mostrados). También se encontró por Pfam que UgpB contiene un segmento transmembranal para UgpB, con una región transmembranal del aa 13-35 y un dominio de unión a soluto extracelular (Fig. 4), lo cual corresponde con los resultados de los distintos programas empleados.

Mediante el programa PHD se predijo la estructura secundaria de la región TM, así como su topología, el resultado mostró que los primeros 18 aa del extremo amino se localizan hacia el interior celular, la región TM presenta una estructura de α -hélice de T19-A35 en UgpB de *B. melitensis* y el resto carboxilo de la proteína se encuentra hacia el exterior de la membrana (Fig 5A), el resultado de la estructura secundaria de la región TM, fue corroborado mediante el servidor PsiPred (Fig. 5B) y Jpred (no mostrado), con ambos programas se determinó la estructura secundaria del precursor de UgpB de T3-T30, es importante señalar que hay variación según el programa empleado, en cuanto al tamaño de la región transmembranal, en aproximadamente 2 aa antes ó después, sin embargo, es necesario tomar en cuenta que las proteínas de estructura conocida por cristalografía de rayos X, son en su mayoría globulares, no se conocen muchas proteínas transmembranales, tan solo una de cada 100 proteínas cristalizadas presentes en el PDB es transmembranal, por eso corroboramos los resultados predictivos mediante varios programas.

operon en *T. litoralis* es muy similar al de *E. coli* y otros sistemas de transporte ABC bacterianos dependiente de proteína de unión, como los genes *ugp* del sistema de transporte ABC del género *Brucella* [33].

A pesar de que 1 de cada 100 proteínas en el banco de datos son proteínas transmembranales y que la homología entre TMBP y UgpB de *Brucella* es del 19 % (identidad de secuencia baja), Swiss-Model fue capaz de generar el modelo para la proteína UgpB de *Brucella* a partir de TMBP. Considerando la similitud en las características bioquímicas de TMBP y las putativas de UgpB de *B. melitensis* así como evidencias experimentales [9], el modelo construido puede ser considerado para explicar posibles funciones de la proteína UgpB.

Swiss-Model generó el modelo tridimensional para UgpB, mostrando dos fragmentos polipeptídicos separados, el primero de T81-C199 y el segundo de G172-A319, dado que no se realizó el modelado de los primeros 80 aa, en adelante se considera solo el nombre de UgpB para referirse al modelo tridimensional propuesto para la proteína de *B. melitensis*, *B. suis* y *B. abortus*. Los fragmentos separados tienen 29 aa en común y con el fin de visualizar mejor la proteína en conjunto, los polipéptidos fueron sobrepuestos con el programa Deep View/Swiss PDB Viewer, mediante alineamiento estructural de los dos polipéptidos como se observa en la Fig 6 A y B.

Después de haber realizado el alineamiento estructural se encontró (Fig 6C) que de los aminoácidos superpuestos, los 8 últimos no pudieron ser alineados, como lo muestra el valor de rms (desviación cuadrática media), identificando a los aminoácidos sobrepuestos más alejados uno de otro con rms mayor a 5 Å en color rojo, estos aa no se pudieron alinear estructuralmente, en cambio, de los primeros 21 aminoácidos superpuestos de ambos fragmentos peptídicos, localizados más cercanos al extremo amino, 15 pudieron alinearse estructuralmente como se aprecia en la Fig 6 C, a los cuales el Deep View/Swiss PDB Viewer les asignó color azul a verde, indicando que la distancia que separa a estos aa alineados es muy pequeña, con rms de 3- 0 Å (rms=0, en azul oscuro) (Fig 6 C), en cambio, de los últimos 8 aa, solo 2 muestran rms cercanos y los valores rms de los otros 6 aa son mayores (indicado en color rojo), por lo que no pueden empalmarse con la estructura de referencia.

3.3.2 Análisis de la homología del sitio de unión a ligando de TMBP con UgpB

La homología entre UgpB y TMBP corresponde al sitio de unión al ligando extracelular, en la figura 7A se muestra la estructura tridimensional de TMBP y en color morado los aminoácidos que participan en la unión a trehalosa.

La tabla de la figura 7C, muestra los aminoácidos de TMBP que participan en la interacción con trehalosa y los aminoácidos homólogos de la proteína UgpB de *Brucella*. Los 26 aminoácidos implicados en la homología no son idénticos, pero algunos muestran relación estructural, perfil hidrofóbico y propiedades fisicoquímicas similares.

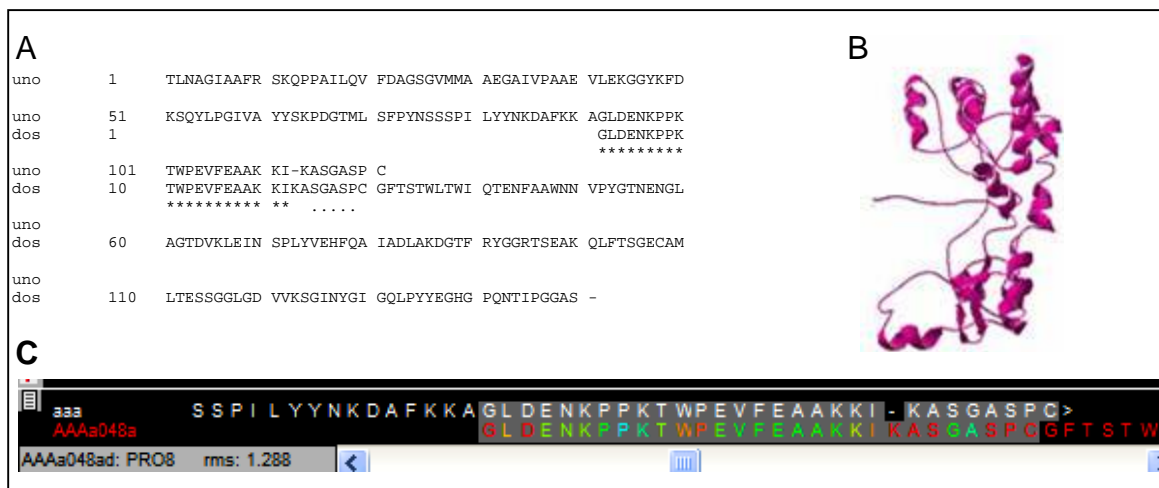


Fig. 6. Sobreposición de los dos fragmentos polipéptidos de UgpB, generados por Swiss-Model. A) Alineamiento de los 2 fragmentos; uno, primer fragmento; dos, segundo fragmento. **B)** Estructuras terciarias alineadas. **C)** Asignación de color al valor de rms (rojo, rms mayor a 5 Å; del verde al azul, rms de 3- 0 Å). La numeración corresponde al modelo tridimensional propuesto y no a la secuencia aminoacídica total de UgpB.

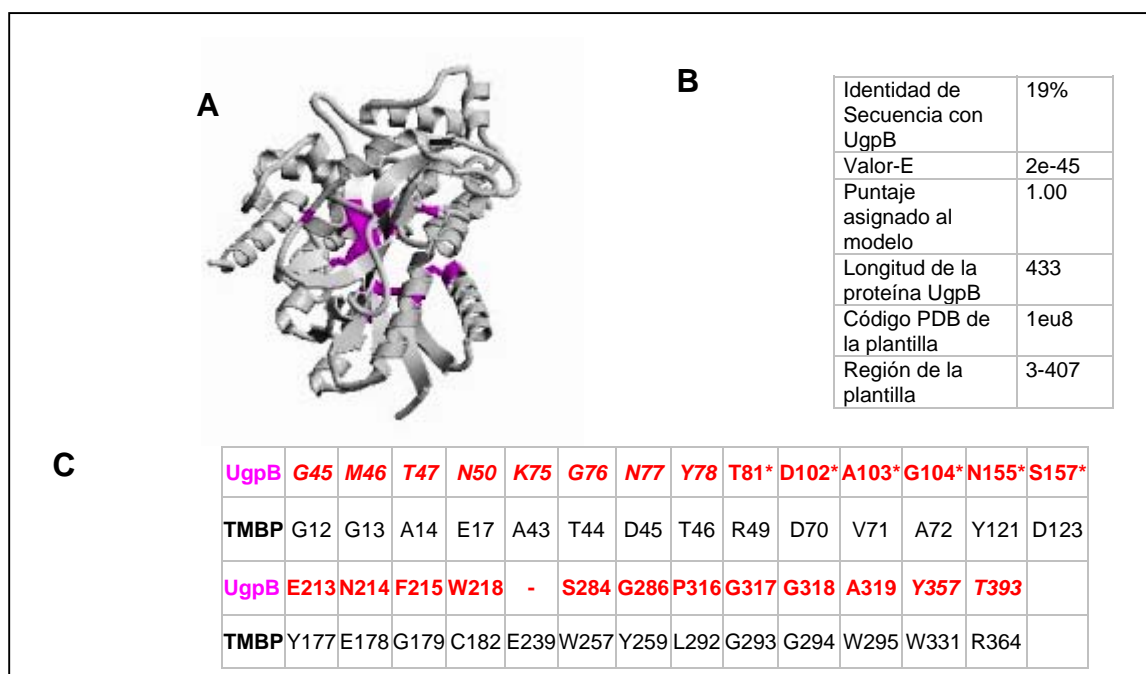


Fig. 7. Homología del sitio de unión a ligando de UgpB con la proteína TMBP. A. Cristal de la proteína de unión a trehalosa/maltosa (TMBP, 1eu8), con el sitio de unión a ligando marcado en color morado. **B.** Datos de la homología de UgpB con el sitio de unión a TMBP **C.** Residuos involucrados en la unión a ligando correspondientes a UgpB (en rojo) y a la proteína plantilla TMBP (en negro).

Ambas estructuras, UgpB modelada por Swiss-Model y la estructura de cristal TMBP uniendo al ligando trehalosa (obtenida del PDB), fueron superpuestas con el programa Deep View/Swiss PDB Viewer.

Después de realizar el alineamiento estructural (por cadenas laterales) y el cálculo por rms se pudo apreciar en el empalme de las estructuras, la superposición de algunas α - hélices, hojas β -plegadas y giros aleatorios (Fig. 8 B), lo cual nos muestra que la homología

estructural entre estas dos proteínas está más conservada que la secuencia aminoacídica.

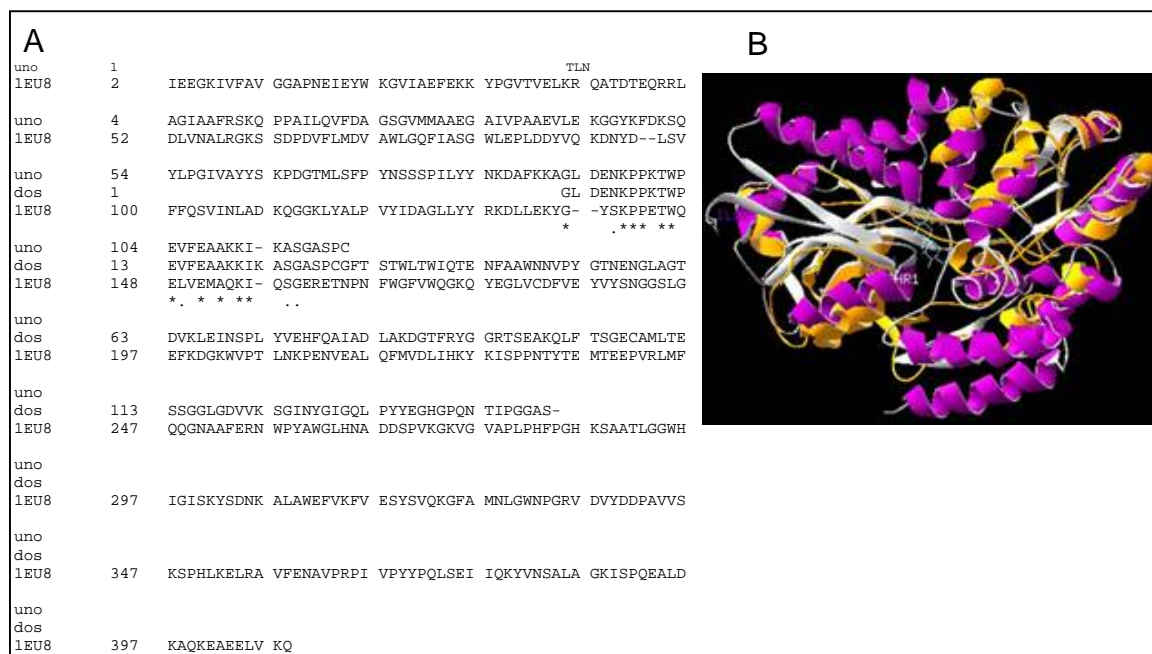


Fig. 8. Sobreposición de las estructuras tridimensionales de UgpB modelada y 1eu8 uniendo al ligando trehalosa. A. Alineamiento Estructural de la proteína de referencia (1eu8) y de los dos fragmentos modelados de UgpB: uno (primer fragmento), dos (segundo fragmento). B. Superposición de TMBP (*T. litoralis*) y UgpB (*Brucella*). TMBP en color morado (hélices) y gris (hojas β-plegadas y giros), al centro y en blanco: trehalosa, UgpB en color naranja.

La interacción de los aminoácidos de la proteína TMBP con el ligando trehalosa se aprecia en la Fig 9 [32]. De los 26 aa que conforman el sitio de unión a TMBP (Fig 7C), 10 interaccionan directamente con trehalosa mediante puentes de hidrógeno e interacciones de Van der Waals y el resto de estos aminoácidos forman parte de la cavidad hidrofóbica. Los aa de UgpB que podrían interaccionar con el azúcar trehalosa, dada su disposición en el espacio y su valor rms (Fig 9A y B), se encuentran predominantemente en el primer fragmento de UgpB modelado por Swiss-Model (Fig 6A), de los cuales 2 forman parte de la cavidad hidrofóbica (A103, G104) y 4 aa (T81, D102, N155 y S157) presentan rms muy cercanos con los aminoácidos R49, D70, Y121 y D123 de la proteína de referencia, los cuales interaccionan directamente con el ligando trehalosa (Fig 10).

La superposición permitió identificar que la disposición en el espacio, de los aa que interaccionan con trehalosa de TMBP (R49, D70, Y121 y D123, color azul, Fig 10A y B), es muy similar a la de los residuos de UgpB (T81, D102, N155 y S157 color morado, Fig 10A y C).

Al visualizar los radios de Van der Waals de UgpB en comparación con los de la estructura de cristal, se aprecia que solo los residuos Asp22 y Ser 77 interaccionan con la trehalosa en las 2 estructuras superpuestas, pero los radios de los residuos Thr1 y Asn 75 de UgpB no alcanzan a interaccionar con los radios de Van der Waals del ligando trehalosa, debido a que sus radicales son más cortos que los correspondientes residuos R49 y Y121 de la estructura de referencia 1eu8.

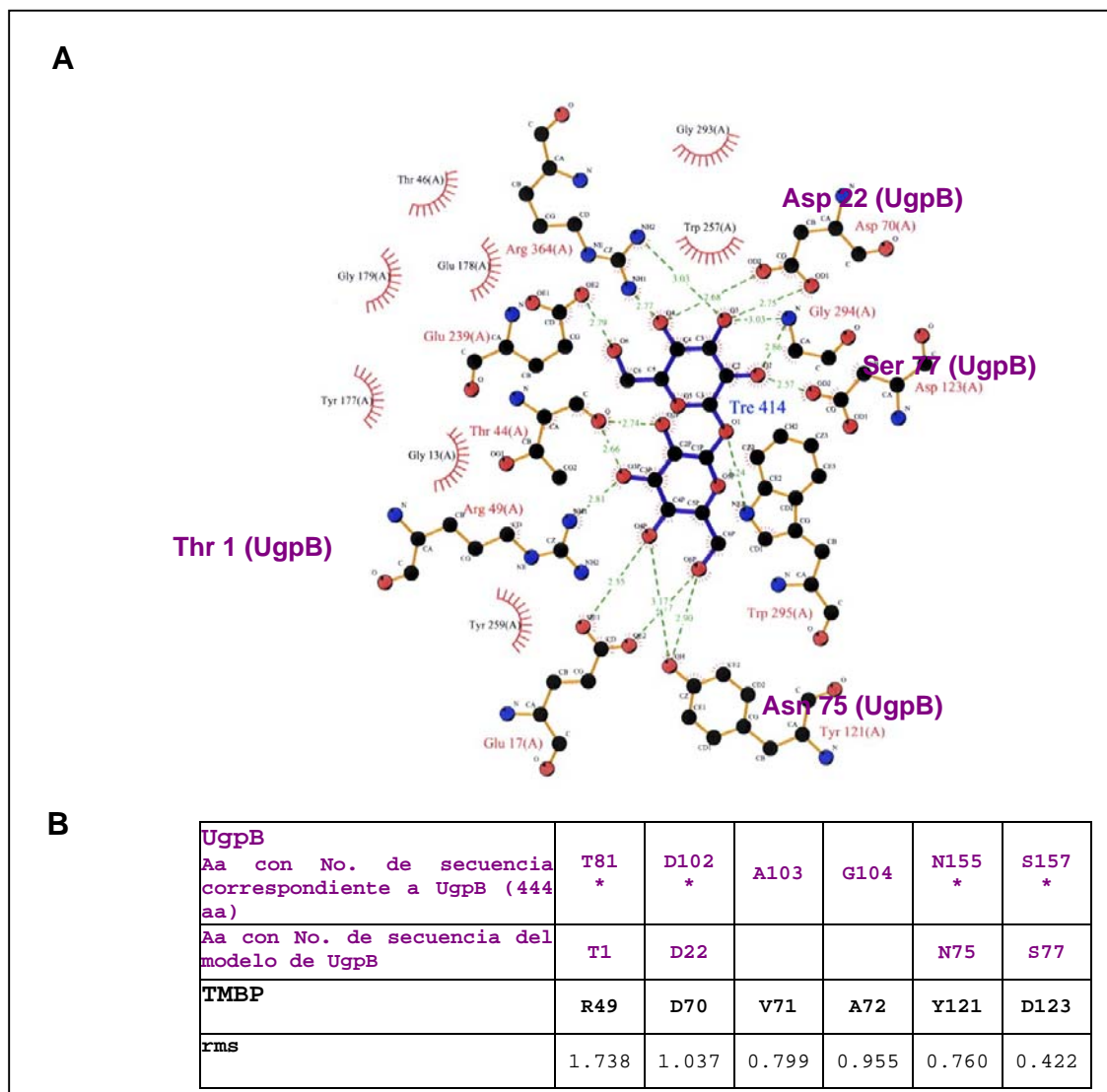


Fig. 9. Sitio de unión a Trehalosa. A. Datos cristalográficos [32] de los aa que interactúan directamente con trehalosa y aa que conforman la cavidad hidrofóbica del sitio de unión.

B. Aminoácidos del primer fragmento de UgpB (*) que presentan rms muy cercanos con los aa de la proteína de referencia TMBP y se superponen con 4 de los 10 aa que interactúan con trehalosa, en color morado aa de la secuencia de UgpB.

La superposición nos sugiere que los residuos Asp22 y Ser77 de UgpB, podrían participar en la interacción con un azúcar, se ha comprobado que el género *Brucella* es capaz de incorporar trehalosa del medio [34], en el hipertermófilo *T. litoralis* se sabe que trehalosa es incorporada para funcionar como un osmoregulador. *Brucella* por su parte, también requiere de osmoreguladores pues puede enfrentarse a medio ambientes salinos como cuando se encuentra en quesos, o aguas de océanos, en donde ya se han reportado evidencias acerca de su presencia en mamíferos marinos [6, 35]. La interacción de trehalosa con UgpB, sería de gran utilidad para que *Brucella* pueda soportar y subsistir en altas concentraciones salinas a las cuales se puede enfrentar en su vida.

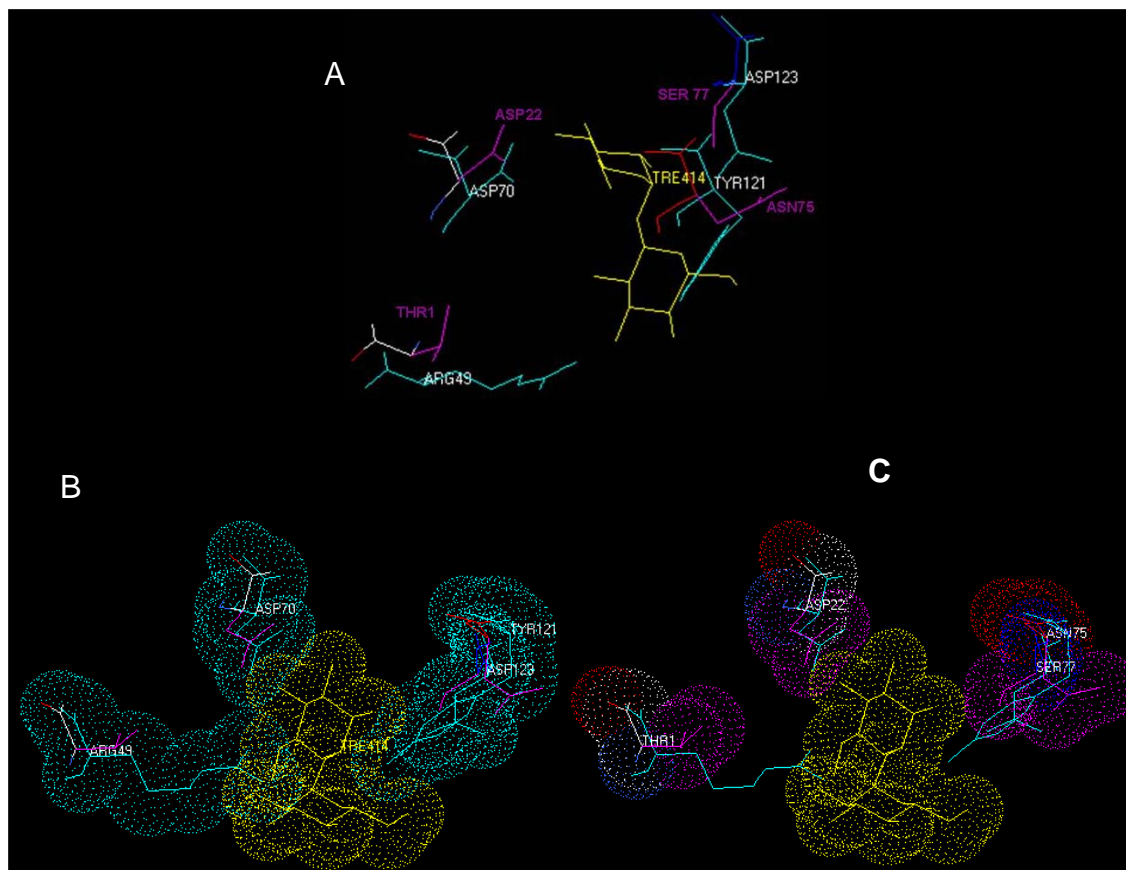


Fig. 10. Aminoácidos que interaccionan con trehalosa. A. Color morado, aa de UgpB involucrados en la interacción con trehalosa. Color azul, residuos aminoacídicos de TMBP involucrados en la interacción con trehalosa. Color amarillo, trehalosa. **B.** Radios de Van der Waals de 1eu8 en color azul y en color amarillo de la trehalosa. **C.** Radios de Van der Waals de de los residuos de UgpB en color morado.

Por otra parte, planteamos la posibilidad de que los residuos Asp22 y Ser77 pudiesen interaccionar con el ácido siálico, azúcar modificado presente en la superficie de diversas células eucariotas en las que se encuentra anclado a proteínas transmembranales. Castañeda et al, en el 2004 demostraron que *Brucella* es capaz de reconocer residuos de ácido siálico presentes sobre la membrana de eritrocitos y de células epiteliales.

Nuestros análisis sugieren que esta interacción pudiera ser mediante los carbonos (C) 8 y 9 que son 2 de los 3 C que forman parte del glicerol, presente en la estructura del ácido siálico (Fig 11A), estos carbonos se encuentran por su conformación espacial (Fig 11B) más accesibles para ser reconocidos por moléculas en el medio exterior (como UgpB de *Brucella*), de tal manera que esta observación nos llevó a pensar en la posibilidad de la interacción de los grupos hidroxilo de los C8 y 9 del ácido siálico con el grupo carbonilo de Asp22 de UgpB.

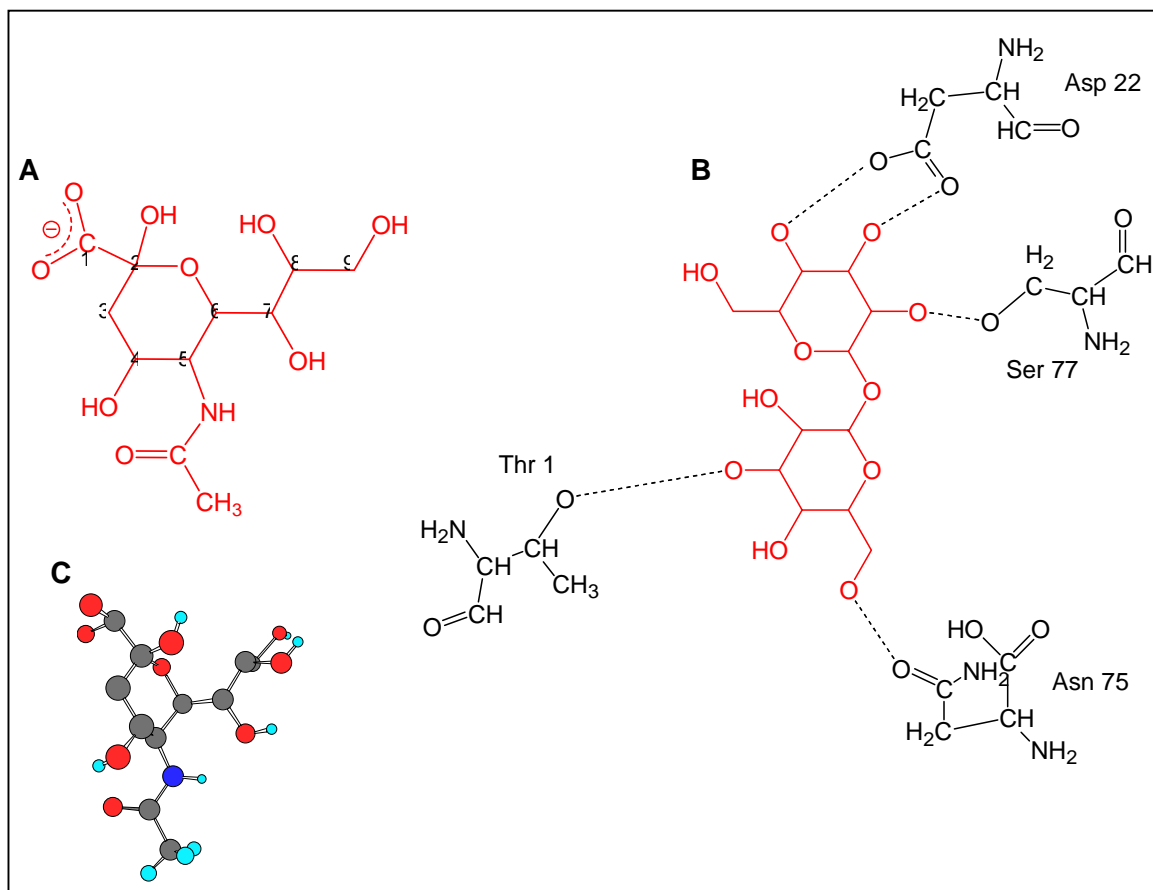


Fig. 11. Posible interacción de trehalosa con UgpB. A. Estructura de ácido siálico. **B.** Asp22 y Ser77 de UgpB superpuestos con la estructura de referencia 1eu8; en color rojo, ácido siálico. **C.** Estructura tridimensional del ácido siálico.

4. Conclusiones

La predicción de la topología y estructura secundaria por diferentes programas, así como de la estructura terciaria y la homología con el sitio de unión de la proteína TMBP, nos sugiere que es muy probable que UgpB de *B. melitensis* sea una proteína transmembranal de tipo II [1], cuyo extremo amino se encuentra localizado al interior de la membrana bacteriana, con un segmento transmembranal, aproximadamente de 20 aa y el extremo carboxilo localizado al exterior de la célula (dominio extracelular), en el cual se localiza el dominio que podría participar en el reconocimiento de estructuras específicas presentes en el medio ambiente de *Brucella*, como son el azúcar trehalosa y el azúcar modificado ácido siálico, lo que le puede permitir a la bacteria, la posibilidad de adaptarse a medio ambientes salinos o identificar receptores de células huésped para iniciar sus mecanismos de adherencia e invasión y producir enfermedad en los organismos de estas células (Fig 12).

La relación estructura-función determinada por predicción, permite tener un panorama más amplio de las funciones de la proteína, lo cual contribuye al mejor entendimiento de los mecanismos desencadenados por los organismos o microorganismos a los que pertenecen dichas proteínas.

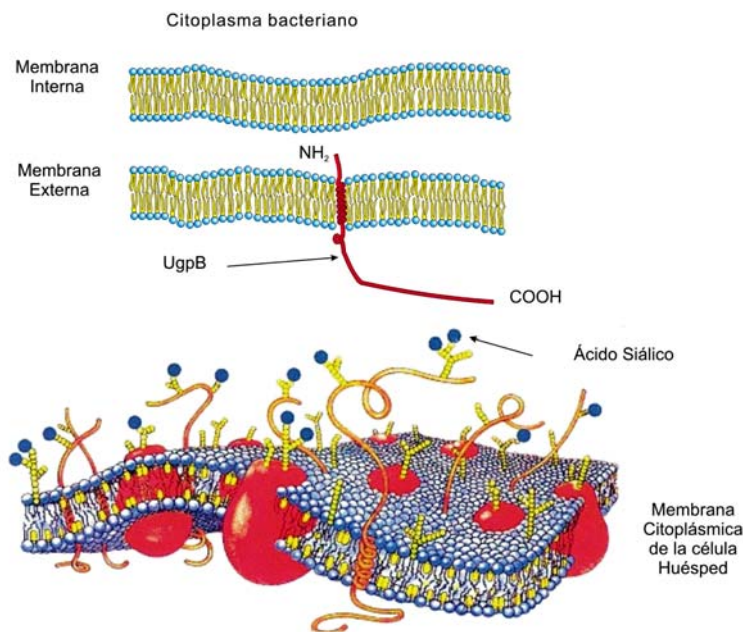


Fig. 12. Modelo propuesto para UgpB de *B. melitensis*. Reconocimiento de ácido siálico de una célula huésped, por medio del locus extracelular de UgpB de *B. melitensis*.

Es necesario identificar los factores que desencadenan la expresión de UgpB, en el género *Brucella* y su relación con los mecanismos de patogenicidad y resistencia a diversos ambientes. El esclarecimiento de los procesos implicados con la expresión, función y regulación de la actividad de UgpB requiere de la interacción interdisciplinaria, de áreas de la ciencia como biología, bioquímica, química, física, matemáticas y ramas multidisciplinarias como la bioinformática. Estos conocimientos abren la posibilidad de diseñar con mayor precisión nanoestructuras que tengan la capacidad de reconocer de manera selectiva, compuestos celulares con el objetivo de inhibir mecanismos patogénicos sin afectar otras vías metabólicas y fisiológicas de las células a tratar.

5. Referencias

- [1] Nelson, D. Michael, Cox. Lehninger Principles of biochemistry. Fourth Edition. Chapter 11. (2004).
- [2] Claverie, J.M., and Notredame, C.: Bioinformatics for Dummies. 1st edn, New York, NY., Wiley Publishing Inc. (2003).
- [3] Pappas G., Papadimitrou, P., Akritidis, N., Christou, L., Tsianos, E. V. The new global map of human brucellosis. *Lancet Infect Dis.* Feb; 6(2): 91-9. (2006).
- [4] WHO. Our planet, our health: Report of the WHO commission on health and environment. Geneva: World Health Organization. (1992).
- [5] Young, E. J. Human brucellosis. *Rev. Infect. Dis.* 5:821-824. (1983).
- [6] Sohn, A.H., Probert W. S., Glaser, C. A., Gupta N., Bollen A. W., Wong J.D., Grace E. M., McDonald W. C. Human neurobrucellosis with intracerebral granuloma caused by a marine mammal *Brucella* spp. *Emerg. Infect. Dis.* 9. 485-488. (2003).
- [7] Godfroid Jacques., Et.al., From the discovery of the Malta fever's agent to the discovery of a marine mammal reservoir, brucellosis has continuously been a re-emerging zoonosis, INRA, EDP Sciences, *Vet. Res.* 36. 313-32. (2005).

- [8] Castaneda-Roldan EI, Avelino-Flores F, Dall'Agnol M, Freer E, Cedillo L, Dornand J, Giron JA. Adherence of Brucella to human epithelial cells and macrophages is mediated by sialic acid residues. *Cell Microbiol.* May;6(5):435-45. (2004).
- [9] Castaneda-Roldan EI, Ouahrani-Bettache S, Saldana Z, Avelino F, Rendon MA, Dornand J, Giron JA. Characterization of SP41, a surface protein of Brucella associated with adherence and invasion of host epithelial cells. *Cell Microbiol.* Jul 4. (2006).
- [10] Kanehisa, M. and Goto, S.: KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28, 27-30 (2000).
- [11] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J.: "Basic local alignment search tool." *J. Mol. Biol.* 215:403-410 (1990).
- [12] Benson A. D., Karsch-Mizrachi L., Lipman J. D., Ostell D. and Wheeler L. GenBank. *Nucleic Acids Research.* 31: 23-27 (2003).
- [13] Higgins D., Thompson J., Gibson T. Thompson J.D., Higgins D.G., Gibson T.J.: CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22: 4673-4680 (1994).
- [14] Falquet L, Pagni M, Bucher P, Hulo N, Sigrist CJ, Hofmann K & Bairoch A.: The PROSITE database, its status in. *Nucleic Acids Res.* 30:235-238 (2002).
- [15] Jackson JD, Ke Z, Lanczycki C, Liebert CA, Liu C, Lu F, Marchler GH, Mullokandov M, Shoemaker BA, Simonyan V, Song JS, Thiessen PA, Yamashita RA, Yin JJ, Zhang D, Bryant SH. CDD: a Conserved Domain Database for protein classification. *Nucleic Acids Res.* Jan 1; 33. D192-6 (2005)
- [16] Marchler-Bauer A, Bryant SH. "CD-Search: protein domain annotations on the fly". *Nucleic Acids Res.* 32 (W) 327-331 (2004).
- [17] Marchler-Bauer A, Anderson JB, Cherukuri PF, DeWeese-Scott C, Geer LY, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Liebert CA, Liu C, Lu F, Marchler GH, Mullokandov M, Shoemaker BA, Simonyan V, Song JS, Thiessen PA, Yamashita RA, Yin JJ, Zhang D, Bryant SH.: "*CDD: a Conserved Domain Database for protein classification.*", *Nucleic Acids Res.* 33: D192-6 (2005).
- [18] Bairoch A., Gastelger E., Gattiker A., Hoogland C., Lachaize C., Mostaguir K., Ivanyi I., and Appel D. R.: The expasy proteome WWW server in 2003. <http://www.expasy.org/doc/expasy.pdf>. (2003).
- [19] Krogh, A., Larsson, B., Von Heijne, G., Sonnhammer, E.L.: Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *Journal of Molecular Biology*, 305(3):567-580, January (2001).
- [20] [Sonnhammer EL](#), [Eddy SR](#), [Birney E](#), [Bateman A](#), [Durbin R](#). Pfam: multiple sequence alignments and HMM-profiles of protein domains. : *Nucleic Acids Res.* Jan 1;26(1):320-2. (1998).
- [21] Sonnhammer, E.L.L. G. von Heijne, and A. Krogh.: A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, 6, 175-182 (1998).
- [22] von Heijne, G.: Membrane Protein Structure Prediction: Hydrophobicity Analysis and the 'Positive Inside' Rule. *J.Mol.Biol.* 225, 487-494 (1992).
- [23] Rost Burkhard .PHD: Predicting one dimensional protein structure by profile-based neural networks. *Methods in enzymology.* 266: 525-539 (1996).
- [24] McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. *Bioinformatics.* 16, 404-405 (2000).
- [25] Schwede T, Kopp J, Guex N, and Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Research* 31, 3381-3385 (2003).
- [26] Guex, N. and Peitsch, M.C. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* 18, 2714-2723 (1997).
- [27] Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Millar W, Lipman DJ.: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389-3402 (1997).
- [28] Boss, W. Binding protein-dependent ABC transport system for glycerol 3-phosphate of *Escherichia coli*. *Methods in Enzymology*, Vol. 292: 4-51 (1998).
- [29] Kyte J., Doolittle R. F. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 157:105-132 (1982).
- [30] Eisenberg D., Schwarz E., Komarony M., Wall R. *J. Mol. Biol.* 179:125-142 (1984).
- [31] Rao M.J.K., Argos P. *Biochim. Biophys. Acta* 869:197-214(1986).

- [32] Diez J, Diederichs K, Greller G, Horlacher R, Boos W, Welte W.J Mol Biol. The crystal structure of a liganded trehalose/maltose-binding protein from the hyperthermophilic Archaeon *Thermococcus litoralis* at 1.85 Å. Jan 26;305(4):905-15) (2001).
- [33] Horlacher, B., Xavier, K., Santos, H. Archaeal Binding Protein-Dependent ABC Transporter: Molecular and Biochemical Analysis of the Trehalose/Maltose Transport System of the Hyperthermophilic Archaeon *Thermococcus litoralis* Journal of Bacteriology, Feb., p. 680–689 Vol. 180, No.3 (1998).
- [34] Essenberg, R.C., Seshadri, R., Nelson, K., Paulsen, I. Sugar metabolism by Brucellae. Veterinary Microbiology 90: 249-261. (2002).
- [35] Clavareau, C., Wellemans, V., Walravens, K., Tryland, M., Verger, J.M., Grayon, M., Cloeckert, A., Letesson, J.J., Godfroid, J. Phenotypic and molecular characterization of a *Brucella* strain isolated from a minke whale (*Balaenoptera acutorostrata*), Microbiology 144: 3267–3273. (1998).

